collection of a large amount of data. This will allow us to regularly assess the health of people, conduct analysis, monitor the dynamics.

For now we can say that there is no service that allows you to get a full integral assessment of your health. This development is unique in its kind and in terms of quality, completeness and functionality it allows to greatly exceed the existing solutions.

We considered the need for monitoring human health and changing the approach to this process. A minimum viable web service product is currently implemented.

Further plans include the introduction of the service as a tool for testing students in the process of physical training, which will also allow debugging and testing the stability of work under heavy load. Then the web service will function in a mode of free access for all comers that will allow to collect and analyze user data, and also to integrate this web service into the work of medical institutions.

DATA SEPARATION FOR TRAINING THE ARTIFICIAL NEURAL NETWORK TO SIMULATE THE SPATIAL DISTRIBUTION OF CHROMIUM IN THE SURFACE LAYER OF THE SOIL

Rakhmatova A.Y.[1,2*], Sergeev A.P.[1,2], Shichkin A.V.[1,2], Buevich A.G.[1,2]

[1)]Ural Federal University, Yekaterinburg, Russia
[2)]Institute of Industrial Ecology UB RAS, Yekaterinburg, Russia
*E-mail: anya_rahmatova@mail.ru

Abstract. An algorithm for dividing data into training and test subsamples to simulate the spatial distribution of chromium in the surface layer of the soil using artificial neural networks (ANN) was proposed. The algorithm takes into account the spatial inhomogeneity of the variable being modelled. The data was obtained during the soil screening on the urbanized area in Novyy Urengoy city. A model, which used controlled separation, had shown more accurate results.

When modelling the prediction of the spatial distribution of the trait, there are two main tasks. First, the choice of a model, which is capable of reproducing the picture of the spatial distribution of the trait with the sufficient accuracy. The accuracy of prediction obtained using ANN is often higher than that of other methods or expert predictions [1]. The ANN model can be applied to measured data obtained during monitoring and used to predict the content of pollutants in uncontrolled places [2, 3].

The second task is to ensure that the chosen model is able to fully realise all its capabilities. The accuracy of the model is largely depends on the correct learning procedure. One of the main way to improve the capabilities of the method is to optimise the procedure for dividing the sample into a training and test subsamples [4]. A random splitting method is commonly used but this approach has a number of weak points. In this work was used Spatial quoting of raw data, which consists of three steps:

1. The survey area was contoured by a convex polygon so that the geodesic line drawn between any two points (real sampling points) was inside this polygon.
2. The polygon was divided into areas (spatial quotas) including the same number of observations. The boundary points of the polygon as well as the maximum and minimum values from each region were included in the training subsample.
3. From each spatial quota, random selection of the points for the training subsample was occurs, so that its share is 70%.

150 samples were taken from South part of Novyy Urengoy (Fig. 1).

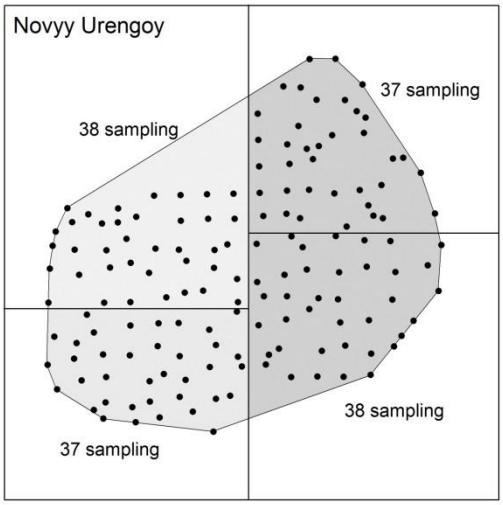

Fig. 3. The survey areas contoured by a convex polygon

The ANN, which used controlled types of splitting turned out to be more accurate than the ANN, which trained by a random splitting for all indicators.

1. Guo G.H., Wu F., Xie F., & Zhang R. Journal of Environmental Sciences, 24(3), 410–418 (2012).
2. Kanevski M. F. International Journal of Systems Research and Information Systems, 8(4), 241–256 (1999).
3. Shaker R. R.,& Ehlinger T. J. International Journal of Applied Geospatial Research, 5(4), 1-20 (2014).
4. Lark R.M,. Marchant B.P. Geoderma, 319, 89–99 (2018).